DOCUMENT RESUME

ED 454 704 FL 026 760

AUTHOR	Kopriva, Rebecca; Saez, Sharon
TITLE	Guide to Scoring LEP Student Responses to Open-Ended Mathematics Items.
INSTITUTION	Council of Chief State School Officers, Washington, DC.
ISBN	ISBN-1-884037-39-9
PUB DATE	1997-00-00
NOTE	43p.; Product of the Council's State Collaborative on Assessment and Student Standards (SCASS), LEP Consortium in collaboration with the Improving America's Schools Act (IASA) Implementation Project.
PUB TYPE	Guides - Non-Classroom (055)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	Elementary Secondary Education; *English (Second Language); Limited English Speaking; *Mathematics Instruction; Performance Based Assessment; Science Instruction; Second Language Instruction; Second Language Learning; Skill Analysis; *Student Evaluation; Teacher Education; *Test Interpretation; Test Theory; Test Validity; Writing Instruction; Writing Skills

ABSTRACT

This training guide is designed to help monolingual English scorers accurately measure the performance of Limited English Proficient (LEP) students on large scale assessment of open-ended mathematics and science performance tests. The guide can also help teachers accurately evaluate classroom work. The guide is arranged in such a way that it can and should be adapted to local training conditions and to needs of various large-scale classroom endeavors. In responding to open-ended assessment items, LEP students are required to read, interpret, devise a solution, write out their mathematical computations, and, very often, communicate their reasoning in writing. Guidance is provided about linguistic issues in the students' native language, developmental benchmarks, associated with learning English, and cultural influences--all of which affect the written responses of LEP students. This document contains linguistic training guidelines, a brief discussion of issues related to the accurate development of assessments for this population, and a glossary of terms. (Contains 12 references.) (KFT)

Reproductions supplied by EDRS are the best that can be made
from the original document.

ED 454 704

Guide to Scoring LEP Student Responses to Open-Ended Mathematics Items

PERMISSION TO REPRODUCE AND DISSEMINATE THIS MATERIAL HAS BEEN GRANTED BY
Cynthia Brown
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
1

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

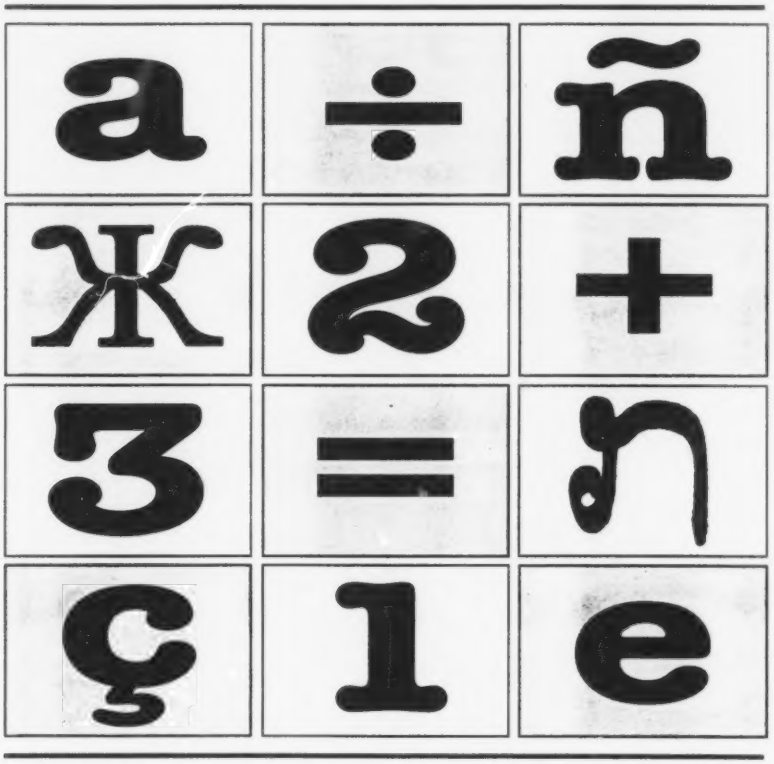
SCASS LEP Consortium Project

Council of Chief State School Officers
1997

BEST COPY AVAILABLE

FLD-6 160

Guide to Scoring LEP Student Responses to Open-Ended Mathematics Items



SCASS LEP Consortium Project

Authors

Rebecca Kopriva and Sharon Saez

Council of Chief State School Officers
1997

Council of Chief State School Officers

The Council of Chief State School Officers (CCSSO) is a nationwide, nonprofit organization composed of the public officials who head the departments of elementary and secondary education in the states, the District of Columbia, the Department of Defense Education Activity, and extra-state jurisdictions. CCSSO seeks its members' consensus on major educational issues and expresses their views to civic and professional organizations, federal agencies, to Congress, and to the public. Through its structure of standing committees and special committees, the Council responds to a broad range of concerns about education and provides leadership on major education issues.

Because the Council respresents the chief education administrators, it has access to the educational and governmental establishment in each state and to the national influence that accompanies this unique position. CCSSO forms coalitions with many other education organizations and is able to provide leadership for a variety of policy concerns that affect elementary and secondary education. Thus, CCSSO members are able to act cooperatively on matters vital to the education of America's young people.

The Council's Assessment Center

The Council's State Education Assessment Center was established to provide an information base on education in the United States, especially from a state perspective. The Center works to improve the breadth, quality, and comparability of data on education, including state-by-state achievement data; descriptive data; indicators of quality in areas such as math and science; and performance assessment of students, teachers, and education leaders. In collaboration with state education agencies, the federal government, and national and international organizations, the Center contributes to a set of useful and valid measures of education geared, when appropriate, to education standards. The Center also supports efforts by states to use standards to improve instruction through collaborative activities of states and others within the education field.

The Council's Resource Center

The Council's Resource Center on Educational Equity was established by chief state school officers to provide services designed to ensure equitable, high-quality and developmentally appropriate education for all students, especially minorities, girls, students with disabilities, limited English proficient students, and low income students. The Resource Center conducts research and policy formulation, develops reports and other materials, operates grant and other action programs, provides capacity building technical assistance to state education agencies, holds working conferences, and monitors federal and state civil rights and education programs that focus on disadvantaged students.

Council of Chief State School Officers

Henry R. Marockie (West Virginia), President

Wilmer S. Cody (Kentucky), President-Elect

Robert E. Bartman (Missouri), Vice President

Gordon M. Ambach, Executive Director

Wayne Martin, Director

State Education Assessment Center

Cynthia G. Brown, Director

Resource Center on Educational Equity

ISBN: # 1-884037-39-9

Copyright © 1997 by the Council of Chief State School Officers, Washington, DC

Acknowledgements

This guide is a product of the Council's State Collaborative on Assessment and Student Standards (SCASS) LEP Consortium in collaboration with the Improving America's Schools Act (IASA) Implementation Project. Support for this project comes from the member states of California, Texas, Delaware, Alaska, Department of Defense Dependents Schools (DODDS), and Connecticut. The consortium activities are coordinated by Council staff members Julia Lara and Edward Roeber. The activities that led to the development of this guide were coordinated by Rebecca Kopriva, formerly with the Delaware SEA and now an independent consultant and Sharon Saez, formerly with the Delaware SEA and now with the U.S. Department of Education. During the initial development of this manual, CCSSO obtained input from colleagues at the National Center for Research in Mathematical Sciences Education (NCMRSE) at the University of Wisconsin, Madison. We thank the following people for their review of this document: Walter Masuda, Michael Ford, Dominick Perissini, Laura Ayala and Mary Ramirez. We specially thank Carolyn Kessler and Joel Gomez for contributing language that greatly enhanced the clarity of the linguistic discussion contained in this text.

TABLE OF CONTENTS

The Council's Assessment Center i

The Council's Resource Center i

Acknowledgements ii

List of Acronyms v

1.0 Introduction 1

1.1 The SCASS LEP Project 1

1.2 Development of the Guide to Scoring LEP Student Responses 1

1.3 Background Information about ELL Students 2

1.3.1 Issues Related to the Development of Language Proficiency 2

1.4 Interpreting Mathematics Assessment Items 3

2.0 Effects of English Language Development on ELL Student Responses 5

2.1 Linguistic Issues 5

2.1.1 Native Language Influences 5

2.1.1.1 Code Switching 6

2.1.1.2 Transposition of Words 7

2.1.1.3 Phonetics 7

2.1.1.3.1 Spelling 8

2.1.2 English Phonetic Influences 8

2.1.3 Merging of Words 14

2.1.4 Omissions 15

2.1.5 Interpretation of Sounds 16

2.1.5.1 Omission of Final Consonant Sounds and Vowels 16

2.1.5.2 Other Examples 16

2.2 Cultural Influences 17

2.2.1 Symbols, Characters, Markings, and Accents 17

2.2.1.1 The Use of Periods Instead of Commas . . . 17

2.2.1.2 Computational Symbols 17

2.2.2 Monetary Systems 17

2.2.3 Metric vs. U.S. Standard Systems 17

2.2.4 Numbers 18

2.2.5 Writing and Mathematics Conventions 18

2.3 Cultural Influences on Stylistic Preferences 19
2.3.1 Circular Responding 19
2.3.2 Long Descriptive Sentences 19
2.3.3 Deductive Reasoning Approach 19
2.3.4 Abbreviated Reasoning Approach 20
2.4 Issues Related to Language Acquisition Development 20
2.4.1 Substitution 21
2.4.2 Confusion in Meaning 21
2.4.3 Use of Unknown Words 21
2.4.4 Novice Sentence and Paragraph Structures 22
2.4.5 Limited Use of Language 22

3.0 Effects of English Language Development on Understanding and Interpreting Assessment Items 23
3.1 Cultural Influences 23
3.1.1 Misunderstanding the Meaning of Words 23
3.1.2 Misunderstanding Systems and Symbols 24
3.1.3 Misunderstanding Because of Differences in What Is Valued and Experienced 24

4.0 Issues Related to the Construction of Accurate Assessments for ELL Students . . 25
4.1 Validity Considerations 25
4.1.1 Item and Test Development 25
4.1.2 Language Complexity 26
4.1.3 Primary Language Assessment 27
4.1.4 Item Overload 27
4.2 Recommendations 27

5.0 Summary 29

Appendix A SCASS LEP Project Trainer's Manual Development Committee . . . 30

Appendix B References 31

Appendix C Glossary 32

List of Acronyms

BICS	Basic Interpersonal Communication Skills
CALP	Cognitive Academic Language Proficiency
CCSSO	Council of Chief State School Officers
DODDS	Department of Defense Dependents Schools
ELL	English Language Learner
IRT	Item response theory
LEP	Limited English Proficient
NCMRSE	National Center for Research in Mathematical Sciences Education
NCTM	National Council of Teachers of Mathematics
SCASS	State Collaborative on Assessment and Student Standards
SEA	State Education Agency

1.0 INTRODUCTION

1.1 The SCASS LEP Project

The State Collaborative on Assessment and Student Standards (SCASS) for Limited English Proficiency (LEP) is one of ten SCASS projects administered by the Council of Chief State School Officers (CCSSO). The project is jointly administered by staff from the Council's State Education Assessment Center and its Resource Center on Educational Equity. The SCASS projects assist states in developing student standards and assessments by working with other states with similar interests. SCASS improves the quality of the student assessments that states are developing and using, reduces the time that it takes for innovations in assessment to be adopted on a wide-scale basis, and reduces the costs required to develop these assessments.

As its overall goal, the Assessing Limited-English Proficient Students SCASS develops procedures and materials for more appropriate assessment of English language learning students, including research on effective programs for English language learning students, language proficiency measures, and other materials related to measuring academic achievement. [1]

1.2 Development of the Guide to Scoring LEP Student Responses

Members of the SCASS LEP Consortium have developed this training guide to help scorers accurately measure the performance of LEP students on large-scale assessment open-ended mathematics and science performance items. The Guide to Scoring LEP Student Responses to Open-Ended Mathematics Items was developed in response to the need identified by member states. It is the product of a series of meetings convened in Washington, D.C., in 1995. (See Appendix A for a list of committee members.)

Teams of mathematics educators from member states (i.e., Connecticut, Delaware, California, and Texas) and others interested in collaborating with CCSSO on this project (i.e., Florida), discussed at length the linguistic features that were apparent in the mathematics responses of LEP students with whom they work. The features specified in this guide were determined to be the most salient features for monolingual scorers to be aware of in order to score LEP students papers. Hence, not all linguistic features that might be shown in LEP students' responses are reflected in the discussion contained here.

Three other products of the SCASS LEP Consortium will be a guide to scoring science open-ended items, a report analyzing the results of

[1] The term LEP and ELL will be used synonymously in this document.

piloting the linguistic training that took place January 1997, and training guidelines. The report will document the results of a study that sought to ascertain the degree to which the scoring of LEP students papers could be improved by training scorers to be sensitive to the linguistic features identified in this guide. The training guidelines will include suggestions of how to integrate the scoring guide into the regular training of scorers and other key personnel. The guidelines will contain suggestions to adequately monitor scorers in large-scale scoring situations, after they have been trained and during the actual scoring process, to ensure that the indicators are used correctly.

This guide identifies indicators that can be used in the training of monolingual English scorers of open-ended mathematics problems from high-volume district or statewide achievement assessments. The information can also be used by teachers to aid them in accurately evaluating classroom work. The guide is arranged in such a way that it can and should be adapted to local training conditions, and to the needs of various large-scale and classroom endeavors. In responding to assessment items written in English, English Language Learners (ELLs) are required to read, interpret, devise a solution, write out their mathematical computations, and, very often, communicate their reasoning in writing.

This document contains linguistic training guidelines, a brief discussion of issues related to the accurate development of assessments for this population, and a glossary of terms.

1.3 Background Information about ELL Students

English Language Learners are students from homes where a language other than English is the principal means of communication. Some were born and raised abroad, while others were born and reared in the United States. The degree to which they were educated and are literate in their native language varies from very little to extensively as does the amount of academic instruction they have received in English before test administration. Consequently, while all ELLs are in the process of acquiring English, they are typically at different stages of acquisition with respect to conversational English (Basic Interpersonal Skills) or BICS and to the use of academic language (Cognitive Academic Language Proficiency) or CALP in the various content areas. Likewise, their English reading and writing levels tend to vary considerably.

1.3.1 Issues Related to the Development of Language Proficiency

Most states require ELLs to participate in districtwide and statewide tests after their second or third year of schooling in the United States and/or after having achieved an intermediate level of English

language proficiency. The belief that level of language proficiency has little or no effect on the learning of mathematics, and therefore, on the assessment of mathematical knowledge, has been proven to be a myth. ELLs who are in the process of acquiring a basic knowledge of English vocabulary, syntax, and semantic properties, will most likely have problems in understanding and interpreting mathematical assessment items that are language based. In addition, students who are in the process of learning English will develop a degree of proficiency in social language skills (survival English) before they develop any degree of academic language skills. Research indicates that the development of precise academic terms and ways to structure academic explanations and arguments usually takes anywhere from four to six years (Collier, 1987).

Between the time when ELLs begin to participate in district and/or statewide assessments, and the time when their mastery of the English language is commensurate to that of their native English-speaking peers, and they can be successful in the academic arena, it becomes crucial to be able to distinguish what students know in a subject area (e.g., mathematics) and how well they can read and interpret what is required to successfully and clearly articulate their responses in English.

Certain standard patterns of misunderstandings can be identified, which should be helpful to scorers as they attempt to evaluate what ELL students know. Examples of those are described in this manual.

1.4 Interpreting Mathematics Assessment Items

ELLs will interpret and respond to mathematics assessment items in English with differing degrees of success depending upon their mathematics background socio-cultural experience and their proficiency in handling the academic language of mathematics in English at the time they take the assessment. Some, because of poor backgrounds in mathematics with or without adequate English proficiency, will respond not at all or incorrectly in most cases. Others who have good mathematics backgrounds and comprehend mathematics assessment items in English, but have limited proficiency for expressing themselves in English, may calculate answers accurately, but may not be able to articulate their reasoning in written form where assessment items require them to do so. Others may articulate their reasoning adequately but in ways that are difficult for the scorers to grasp, especially under the pressure of a rapid, high-volume scoring situation.

Some students, either because of their stage of English language proficiency at the time of assessment administration, their education abroad, or their cultural background, may interpret the items and problems in ways that may be unexpected or that appear atypical for a scorer. Such background characteristics may prompt students to focus on certain types of information over others or to emphasize aspects of a problem or its

solution that would not be accorded a similar priority by the scorer. In any case, although the scorer may perceive certain interpretations of the meaning of the item, the responses of some ELL students may nevertheless be understandable and defensible, and the solution presented may be acceptable and consistent with the interpretation and the mathematics required by the assessment developers.

It is important for scorers to understand what ELL students are being asked to demonstrate on a math performance-based assessment item. Because it is typical for on-demand assessments to require written responses, the information we present in this manual gives readers information about the linguistic issues that can be confounding factors in assessing an ELL's responses. ELL students are asked to demonstrate not only math skills, but also reading and writing skills in a language that they have not yet fully acquired. This poses a challenge to both the student and the individual who scores the items since they must differentiate between evaluating the student's knowledge of math with accuracy despite the hurdles of functioning in a second language.

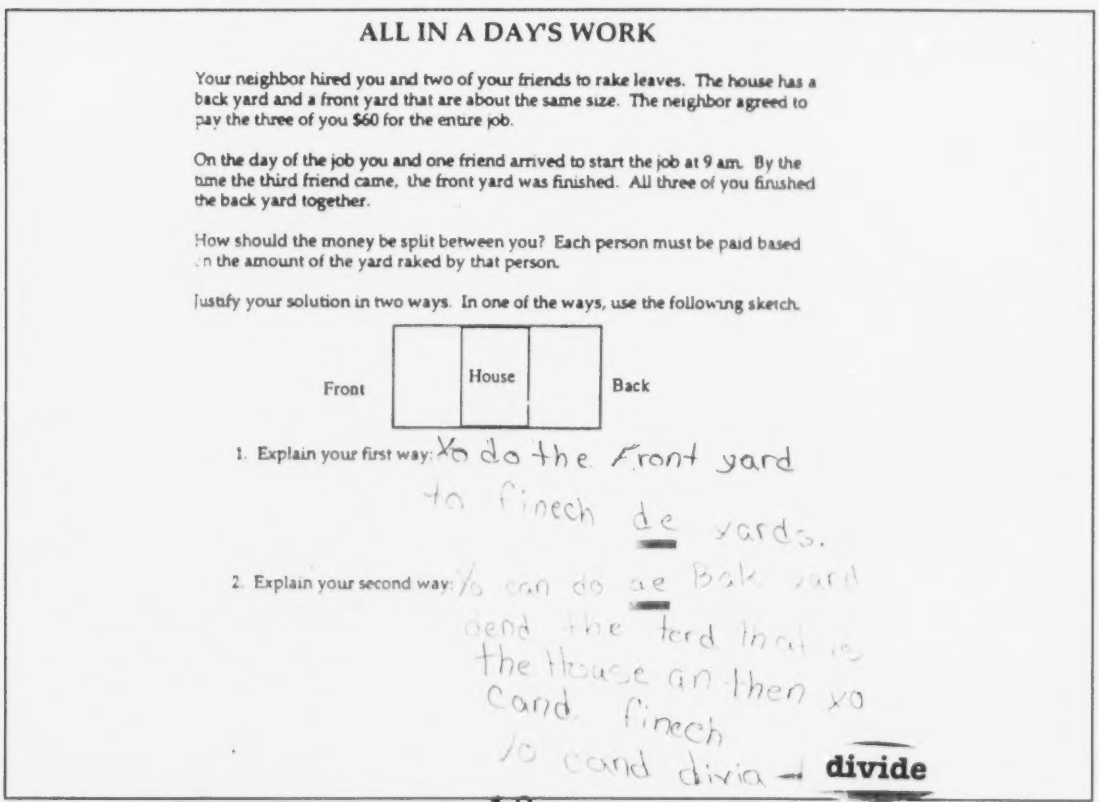
2.0 EFFECTS OF ENGLISH LANGUAGE DEVELOPMENT ON ELL STUDENT RESPONSES

2.1 Linguistic Issues

Linguistics is the science concerned with a number of systems that constitute language and communication. As such, linguists study language at the level of sounds (phonology), words (morphology), sentences (syntax), meaning (semantics), and use (pragmatics). Student responses to test items are products of these five systems working with the English writing system. They are also the products of the interaction between prior knowledge or experience and new knowledge. In the case of second language learners, the process is complicated in the sense that there are more linguistic resources at their disposal—the primary language and the possibility of multiple writing systems from which the product (the test responses) may be constructed (Ford, 1996).

2.1.1 Native Language Influences

Native language influences may appear in ELL student responses. For example, in figure 2.1 the student uses the primary language "d" dental sound in place of the "th" sound in the word *the*.



ALL IN A DAY'S WORK

Your neighbor hired you and two of your friends to rake leaves. The house has a back yard and a front yard that are about the same size. The neighbor agreed to pay the three of you $60 for the entire job.

On the day of the job you and one friend arrived to start the job at 9 am. By the time the third friend came, the front yard was finished. All three of you finished the back yard together.

How should the money be split between you? Each person must be paid based on the amount of the yard raked by that person.

Justify your solution in two ways. In one of the ways, use the following sketch.

Front | House | Back

1. Explain your first way: Yo do the Front yard to finech de yards.

2. Explain your second way: Yo can do ae Bak yard dend the ford that is the House an then yo cand finech yo cand divia → divide

Figure 2.1. Spanish

BEST COPY AVAILABLE

2.1.1.1 Code switching

Code switching or code mixing of languages may appear in the oral or written samples of the second language user. Code switching is the alternate use of language where a sentence may contain elements from both the first language and the second language. Code switching may appear within and/or between sentence structures at the single word, phrase, clause, or sentence level. These influences should not detract from the essential message being conveyed by the respondent.

- Code switch at the word level

  "I think sung *es* becuse is big and is nice."

  (I think the sum is because it is big and nice.)

- Code switch at phrase level

  "I put the forks *en las mesas.*"

Figure 2.2. shows code switching within a sentence at the single word level.

The Vending Machine

Maria wants to buy a 75-cent snack from a vending machine. The machine takes only nickels, dimes, and quarters. Maria has 7 nickels, 5 dimes, and 2 quarters.

Part 1

Show all of the different ways she could pay for the snack. You may use words, diagrams, or charts.

Minh có thể lấy 2 quarters và 25¢ from chỗ 5 dimes.

WE MAY TAKE TWO QUARTER OR 5 DIMES

2 quarters = 50¢ = 50

5 dimes = 50¢ − 25¢ = 25 / 75¢

5 dimes = 50¢ = 50¢

2 quarter = 50¢ = 25 / 75¢

Figure 2.2. Vietnamese

BEST COPY AVAILABLE

2.1.1.2 Transposition of Words

In addition to code switching, the second language learner may follow the rules of syntax or word order used in the home language. This occurs when students use knowledge of sentence structure from the home language as in the following examples:

Adjective/Noun

"La casa azul"

Transliteration: "The house blue"

Translation: "The blue house"

Pronoun/Verb

"Allez vous"

Transliteration: "go you"

Translation: "Let's go"

2.1.1.3 Phonetics

Students bring together two things as they produce their test responses: a writing system and a linguistic system. The writing system in English is alphabetic, which means that a sound can be represented by a number of symbols. The English language has 38 phonemes and only 26 letters, which means that one letter can represent more than one sound.

It is not uncommon for some second language learners to use sounds from their native language while learning to differentiate between the sound systems of their native language and the unfamiliar sounds of their new language of English. Aside from appearing in their oral expression, these native language phonetic forms are also exhibited in written samples, as shown in figure 2.3.

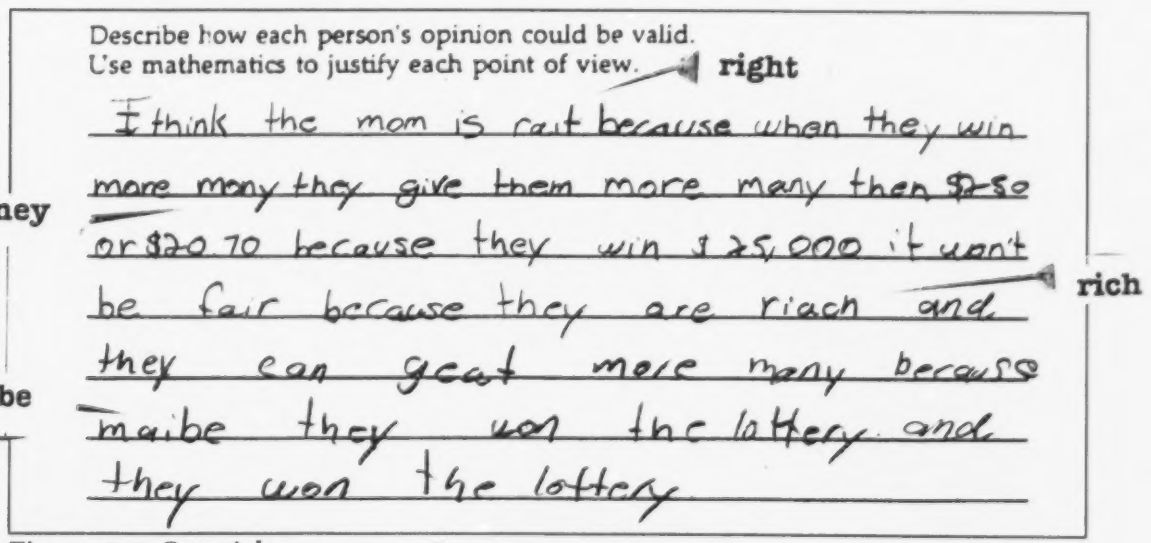

Describe how each person's opinion could be valid.
Use mathematics to justify each point of view.
right
I think the mom is rait because when they win more mony they give them more many then $2.50 or $20.70 because they win $25,000 it wont be fair because they are riach and they can geat more many because maibe they won the lottery and they won the lottery
money
rich
maybe

Figure 2.3. Spanish

BEST COPY AVAILABLE

2.1.1.3.1 Spelling

Students will sometimes use spelling conventions from their first language to write English words. For instance, in Spanish words beginning with "s" and followed by a consonant may be written as "es":

"eschool" or "escul" (school)

Figure 2.4 shows other examples of first language spelling conventions.

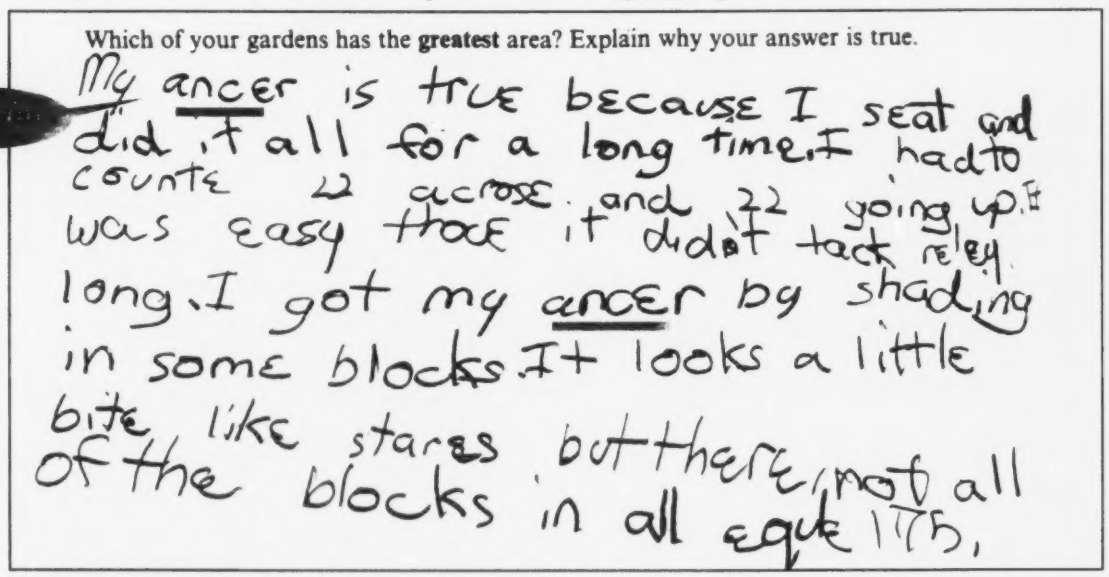

Which of your gardens has the greatest area? Explain why your answer is true.
My ancer is true because I seat and did it all for a long time. I had to counte 22 across and 22 going up. It was easy thoe it didn't tack reley long. I got my ancer by shading in some blocks. It looks a little bite like stares but there not all of the blocks in all equl ITS.

Figure 2.4. Spanish

2.1.2 English Phonetic Influences

ELL students also invent the spelling of English words, given their best estimate from what they know about phonetics in the English language. This is the same process that native English speakers go through when they are learning to write. It is developmentally appropriate, given ELL students' years of experience with English, though often not grade level appropriate in relation to native English speakers.

- Phonetic transposition of letters

  "dose" (does)

  "tow" (two)

  "frist" (first)

  "cions" (coins)

- Phonetic substitution of sounds "b" and "v", "j" for "y", "sh" for "ch", "d" for "th", "ed" for "t"

  "The *pechure* was don *bery* good by the student."

  "The one that had used the least coins is the *lased*"

BEST COPY AVAILABLE

The above examples also show that students are transferring what they know about the phonetics of their native language to write in English. For example, native speakers of Spanish will transpose "b" and "v" in English because in Spanish the "v" is pronounced somewhat like an English "b"—but not quite. A similar situation exists for the "d" and "th." Thus, the above features reflect what students know about phonetics in English and language transfer from first language to English (Masuda, 1996). The following figures are additional examples of phonetic influences (2.5 to 2.8).

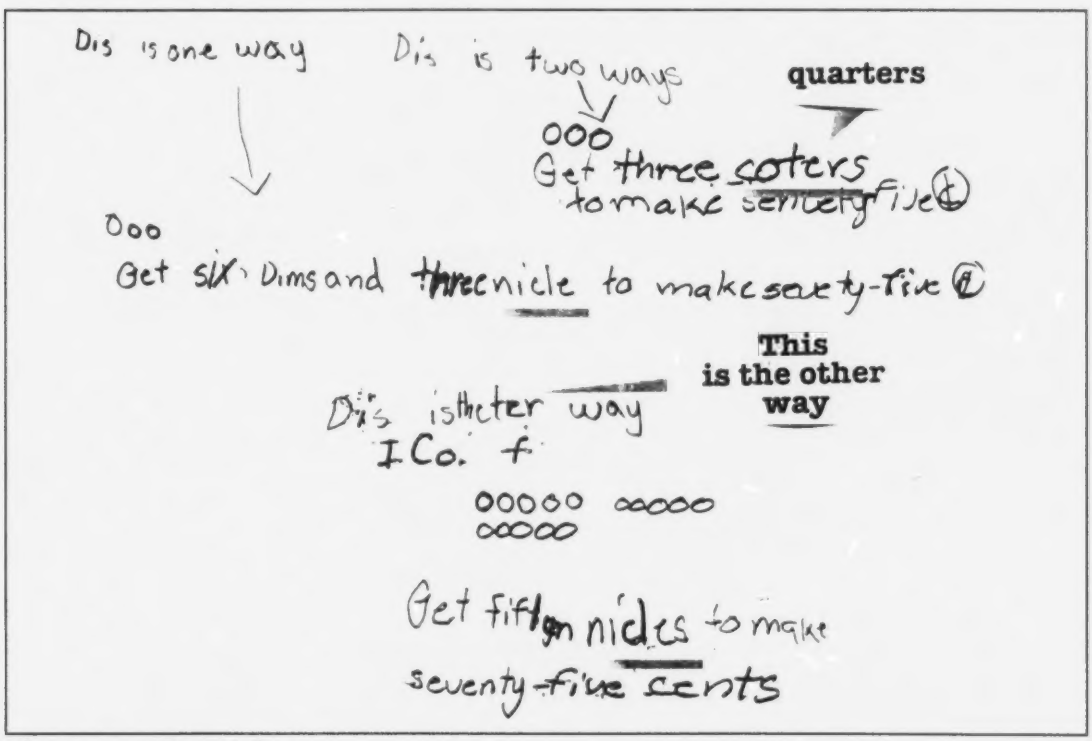

Dis is one way
Dis is two ways
quarters
Get three soters to make seventy five
Get six Dims and threcnicle to make seventy-five
This is the other way
Dis istheter way
I Co. f
Get fifteen nides to make seventy-five cents

Figure 2.5. Spanish

BEST COPY AVAILABLE

Which of your ways uses the **fewest** number of coins? Explain why this is true.

5 dimes and 1 quarter has the lest numbr. Because the rest of them. have almost. All of the mony in the story. That has the few est becaus I only yust 50 and one quarter. And the other one 35¢ or 50 ¢.

Figure 2.6. French Creole

Which of your ways uses the **fewest** number of coins? Explain why this is true.

I put dam fa gather and I add dam

Figure 2.7. Vietnamese

Which of your gardens has the **greatest** area? Explain why your answer is true.

my dus mot he sm but i dot no why my dus mot hasm and its nelu.

Figure 2.8. Spanish **does — have them — like you do**

- Other examples of inventive spelling commonly seen in ELLs responses include those below and in figures 2.9 to 2.15.

  "wat" (what)

  "ecwoles" (equals)

  "yous" or "youd" (use)

  "plast" (plus)

  "wen" (when)

  "in" or "n" (and)

BEST COPY AVAILABLE

Which of your ways uses the **fewest** number of coins? Explain why this is true.

1. I youd two 25¢ i two 10¢ and a 5¢ for number
2. I youd five 10¢ and a 25¢ for number 2
3. I youd treen 5¢ and one 10¢ and two 25¢ for nu

used

Figure 2.9. Spanish

answer a Becuse I only used 7 coins and I used more than other problems

the other way

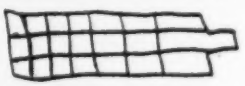
Figure 2.10. Spanish

coins coins

I because it has 8 cions it has 7 cion and 2 has 10 cions and 3 has 11

Figure 2.11. Vietnamese

Which of your gardens has the **greatest** area? Explain why your answer is true.



this one is the largest ereya - area
because it tacks up, takes
the most spas. space

Figure 2.12. Vietnamese

BEST COPY AVAILABLE

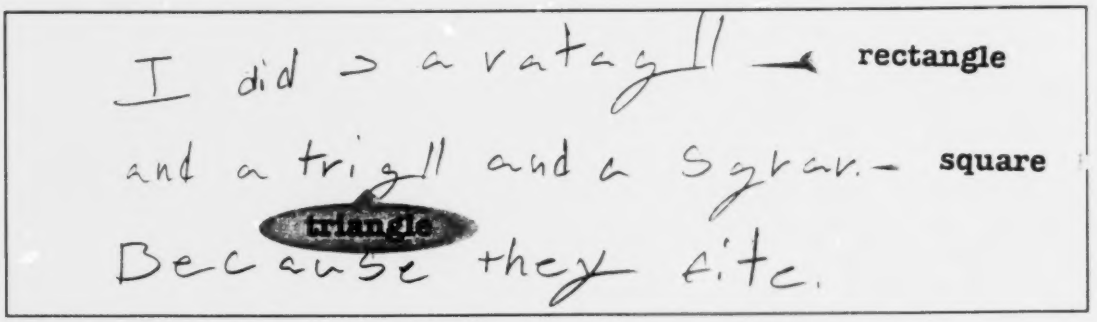


Figure 2.13. Vietnamese

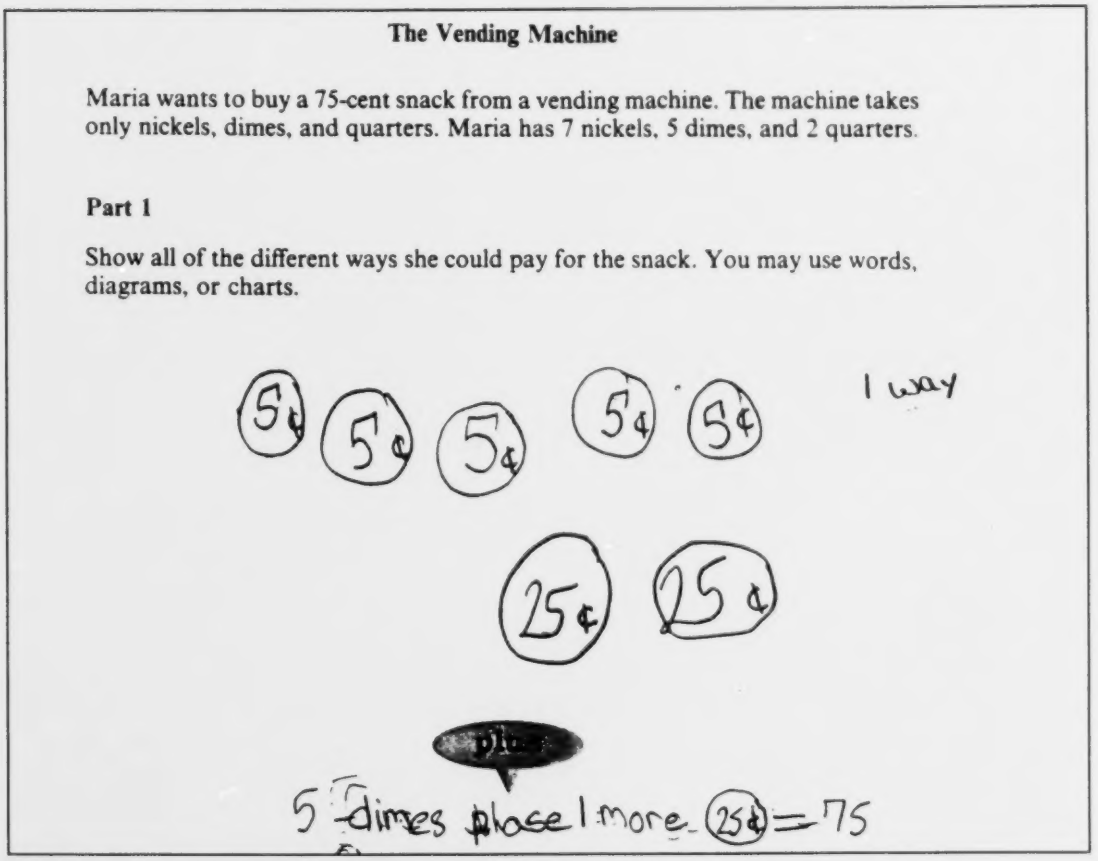

The Vending Machine
Maria wants to buy a 75-cent snack from a vending machine. The machine takes only nickels, dimes, and quarters. Maria has 7 nickels, 5 dimes, and 2 quarters.
Part 1
Show all of the different ways she could pay for the snack. You may use words, diagrams, or charts.
5¢ 5¢ 5¢ 5¢ 5¢
1 way
25¢ 25¢
pl
5 dimes plase 1 more (25¢) = 75

Figure 2.14. Spanish

The Vending Machine

Maria wants to buy a 75-cent snack from a vending machine. The machine takes only nickels, dimes, and quarters. Maria has 7 nickels, 5 dimes, and 2 quarters.

Part 1

Show all of the different ways she could pay for the snack. You may use words, diagrams, or charts. you can add up all the quarters

money - mony ~~2~~quarters and 5 dimes it equs 75 cent and all of the mony equs $1.10

quarters

how much she has

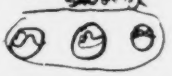


Figure 2.15. Spanish

BEST COPY AVAILABLE

2.1.3 Merging of Words

Closely related to innovative, phonetic spelling is the abbreviation of words and the condensing into one mega-word. Often, second language learners will create a phonetic word form based on their "understanding" of the spoken words. Specifically, second language learners will treat expressions with more than one word as a single word. This may entail phonetically re-creating and grouping various words as a single word. Difficulties may arise when the native language and English language use different sounds, or use sounds that occur occasionally in the English language but pass unnoticed because they are not phonemic. In figure 2.16 the student merges the words *"add"* and *"them."*

"ghuadayamean" (what do you mean?)

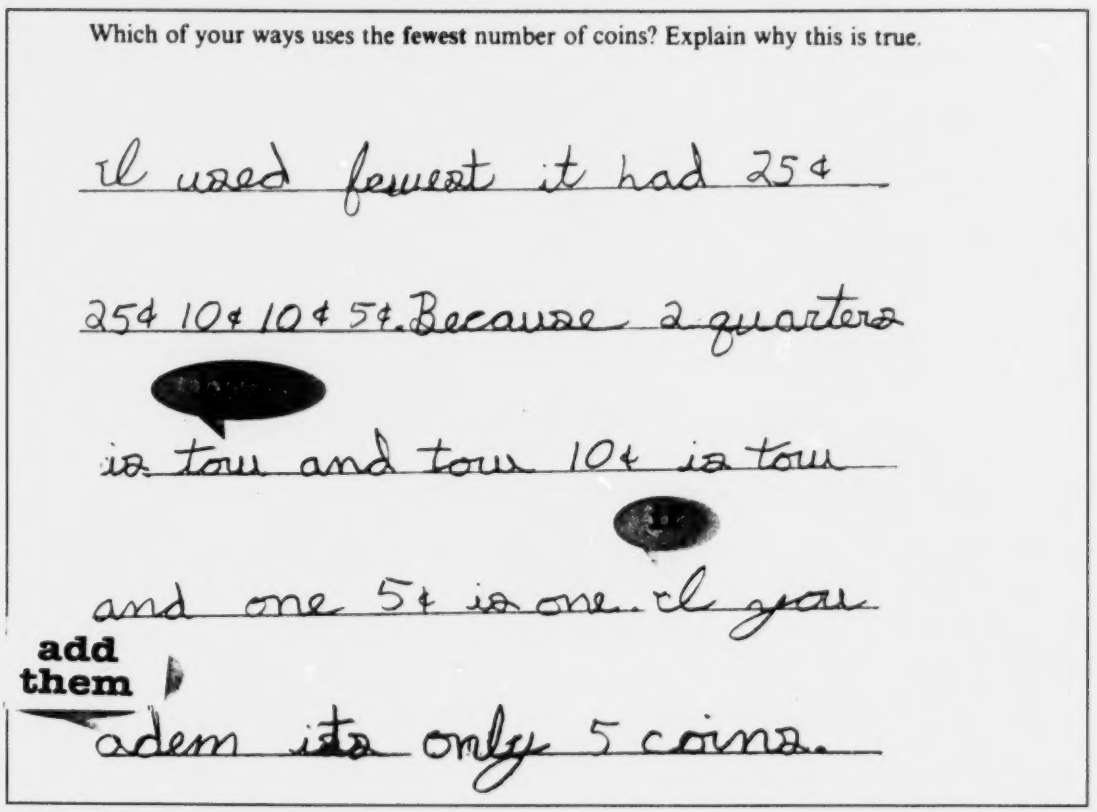

Which of your ways uses the fewest number of coins? Explain why this is true.
il used fewest it had 25¢
25¢ 10¢ 10¢ 5¢. Because 2 quarters
is tow and tow 10¢ is tow
and one 5¢ is one. il yes
add
them
adem ita only 5 coins.

Figure 2.16. Spanish

22

2.1.4 Omissions

Omission of tense markers, articles, plurals, prepositions, or other words in students' written responses can be attributed to many sources, and are seen in the responses of English dominant students. In the responses of second language learners, these omissions may occur because of a lack of understanding of English conventions or because there is no equivalent convention in the student's native language. In figure 2.17, the student omits words and punctuation marks.

"It have [a]fence around it and I counted it."

"Tomorrow I go [to the]bank."

"I walk [walked]."

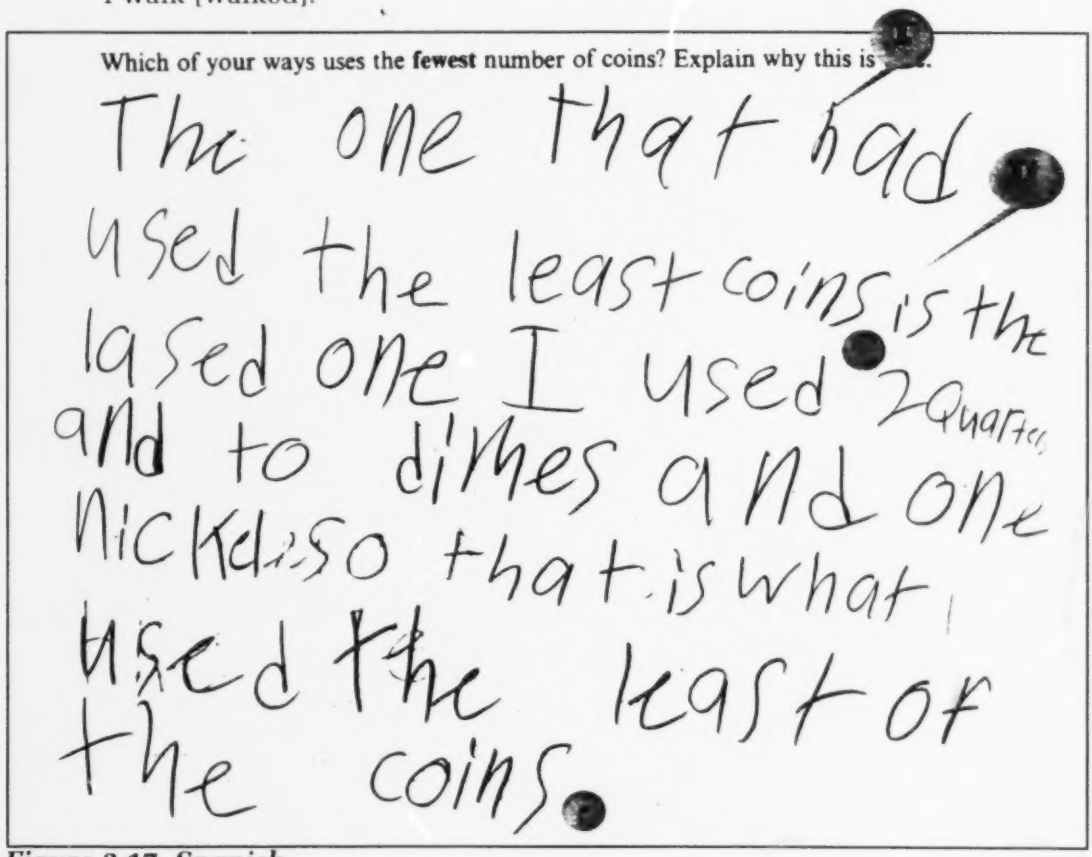

Which of your ways uses the fewest number of coins? Explain why this is
The one that had
used the least coins is the
lased one I used 2 Quarte
and to dimes and one
nickel so that is what
used the least of
the coins.

Figure 2.17. Spanish

2.1.5 Interpretation of Sounds

How ELL students write may be influenced by sounds in their native language or dialect that differ from English sounds. For example, in speaking or writing English, ELL students may omit some final consonant sounds, transpose certain sounds, substitute one sound for another, or reduce consonant clusters.

2.1.5.1 Omission of Final Consonant Sounds and Vowels

"boo" (book)

"wri" (write)

"fatha" (father)

"teacha" (teacher)

"skoo" (school)

The example in figure 2.18 shows the omission of the final consonant.

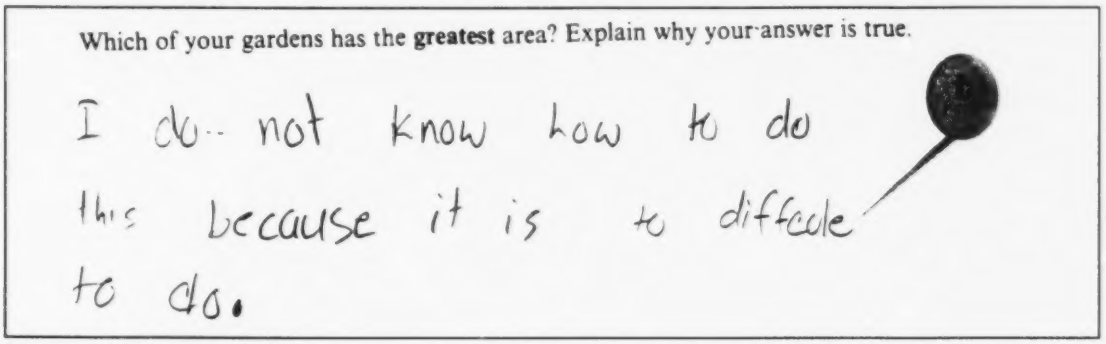

Which of your gardens has the greatest area? Explain why your answer is true.
I do not know how to do this because it is to diffcole to do.

Figure 2.18. Spanish

2.1.5.2 Other Examples

Transposition: "aks" (ask)

Substitution: "piksburg" (Pittsburgh)

Reduction: "picher" (picture)

In some dialects of English, it is entirely appropriate for students to pronounce "ask" as "aks." The same may be true for "piksburg" (Masuda, 1996).

2.2 Cultural Influences

The following points can be considered to be validity issues; that is, they should have been caught in bias review sessions during item development. However, as assessment developers are becoming more aware of the influence of cultures, it is possible that items will be included which contain one or more of the points listed below. We have included the cases in this section because knowledge of these points could help the scorer understand a student response on mathematics items, and subsequently, score the paper more accurately.

2.2.1 Symbols, Characters, Markings, and Accents

Some symbols used in mathematics vary across cultures.

2.2.1.1 The Use of Periods Instead of Commas

Some European and Latin countries use periods instead of commas. For example, a student will read the number 3.001 as three thousand and one, rather than three and one thousandth. Responses from items that contain such numbers should allow for this type of interpretation.

2.2.1.2 Computational Symbols

The following are some examples of the types of symbols that might be confusing to readers.

- In some countries the symbol for long division is written as 3/927.
- In England the "3.5" is often written with the decimal point in a higher position 3·5 (midpoint between two numbers not at the bottom). This could be confused with the symbol for "dot product" in the American mathematical notation for multiplication (Masuda, 1996).

2.2.2 Monetary Systems

There is often confusion related to different monetary systems. Use of monetary words, such as *dollar* or *peso*, may mean different amounts depending on the country.

2.2.3 Metric vs. U.S. Standard Systems

Scorers should be familiar with both systems and be able to evaluate responses from either one (except when the item specifies the use of one system). For instance,

- Distance and volume: miles and gallons vs. meters and liters
- Temperature: Celsius vs. Fahrenheit

2.2.4 Numbers

Numeric symbols are not universal. When the same numeric symbols are used, position, spacing, and simple interpretation of the symbol, expression, and conventions of symbols and certain applications may vary widely.

- In some Latin American countries, a "billion" is read as 1,000,000,000,000, rather than calling it a "trillion."
- Sensitivity to the use of non-Arabic numerals. For example Laotian numbers are represented as follows:

1 2 3 4 5 6 7 8 9 10

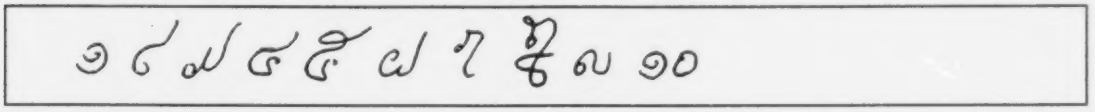


Figure 2.19. Laotian

2.2.5 Writing and Mathematics Conventions

In some countries, the native language is written from right to left, or bottom to top. It is not uncommon to see responses from ELL students where English words or symbols are written accordingly. Sometimes, print or symbols are found written in a circular fashion, or outer to inner rings or the reverse. Also, there are differences in conventions related to mathematics procedures. For instance, the procedure of long division is often written from the item up, rather than placing the numbers below the item as is done in the United States. Figure 2.20 shows an example of how an algorithm is applied in a division problem:

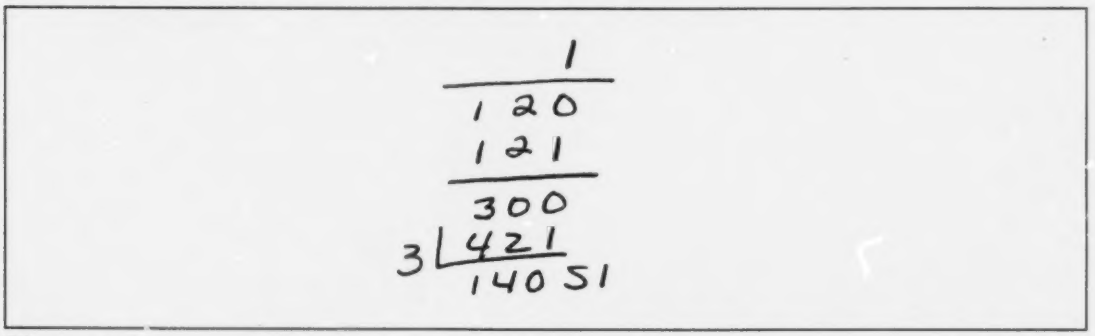


Figure 2.20.

26

2.3 Cultural Influences on Stylistic Preferences

Nonstandard stylistic preferences are separated from other cultural issues, because, while one or more may be prevalent systematically within a culture, this is an issue for all students. Traditionally, assessments have tended to favor a limited range of response and writing styles. The following are some of the more common types of response preferences scorers are apt to encounter among ELL student responses. It should be noted that stylistic preferences may not be evident because of students' lack of proficiency in the language. Their response may simply be a basic verbatim translation with very little understanding of structure, style, or voice.

2.3.1 Circular Responding

A circular style is defined as a style where students appear to be talking around the primary issue. Often they discuss and explain other, less direct, influences to the problem, sometimes including discussions about topics that do not appear initially to be germane to the subject at hand. Eventually these students explain many of the connections and do deliver a response. It is not uncommon for these responses to be fuller and richer than traditional responses. However, they often are wordy, and include some discussions of material the students never link directly to the problem or their response. Often, some components of their understanding, for instance, the mathematical computation, do not show up until fairly late in the explanation.

2.3.2 Long Descriptive Sentences

Closely related to circular responding, and often a component of a circular response, is the use of long, descriptive sentences. Some students embellish their responses more than others, sometimes to a point of frustration for scorers who are in high volume, time constraining situations. Even though scorers are probably sensitized to this issue, it is quite possible that time pressure, or scorer response preference may lead to the assignment of lower scores than is appropriate.

2.3.3 Deductive Reasoning Approach

A person's response style takes the approach of leading up to a point or topic sentence by presenting arguments in a series of often lengthy paragraphs rather than stating the point at the outset. A response employing this style can be scored lower because scorers assume the students are not focused, have less understanding of the problem than they may in fact have, or are simply off topic.

2.3.4 Abbreviated Reasoning Approach

Conversely, because of some students' native language discourse structure and independent system of visual communication (e.g., characters or syllabaries), they prefer a compact, abbreviated response style where every sentence in a paragraph is a topic sentence. Scorers in high-volume situations, who often do not read every word carefully, may miss some important arguments that are presented by students using this style. These responses may be scored lower than is appropriate.

2.4 Issues Related to Language Acquisition Development

Most students first develop a certain level of social communicative competency in the second language (i.e., English). Development and the appropriate use of academic terms and ways to structure academic explanations and arguments usually occur later. A transition period is often noticeable when students have attained a sufficient level of social English proficiency (commonly referred to as basic interpersonal communication skills or BICS). These students are often placed in classrooms where English is only or predominantly used in instruction, and where students are expected to interact, and take tests, in English. Although this stage of second language acquisition is developmentally appropriate, given the ELL students' years of experience with English, it is often not grade-level appropriate in relation to native English speakers. These students may be placed in regular classrooms because their BICS seems fully proficient, even though academic language (CALP) may not be fully developed. This would place the student at a disadvantage if instruction and assessment were carried out with the assumption that the student's academic language were fully developed (or as developed as BICS).

The social, or nonacademic, vocabulary of these students' English language proficiency is still limited relative to the proficiency of peers who are native English speakers. It therefore becomes important to separate what students know in a subject area and how well they can read from what is required to successfully and clearly articulate their responses in English. Certain rather standard patterns of misunderstandings can be identified by the scorers, which should be helpful to scorers as they attempt to evaluate what ELL students know. The following examples relate to scoring mathematical responses.

2.4.1 Substitution

ELLS sometimes substitute common words for precise mathematical terms and concepts, for instance, the use of *fattest* for *greatest*, *smallest* for *fewest*, or *plust* for *added*. Depending on what the item is supposed to measure, the student's substitution may or may not be acceptable. Figure 2.21. is an example of substitution of the word *"fattest"* for *"greatest."*

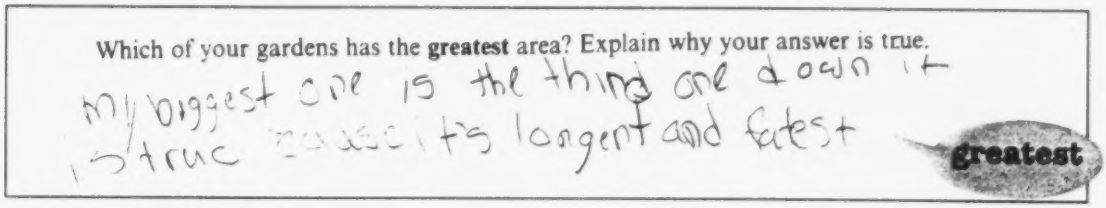

Which of your gardens has the greatest area? Explain why your answer is true.
My biggest one is the thing one down it
struc
greatest

Figure 2.21. Spanish

2.4.2 Confusion in Meaning

Sometimes there is misunderstanding about the meaning of words because of typographical conventions unique to writing in the mathematical language. The confusion often results from the dual meaning of terms used in mathematics and in their own native language, for example, *left* (as opposed to right) vs. *left* (remaining); *whole* (all of the parts) vs. a *whole* number (not a fraction); and *sum* (read as some) vs. *sum* (the result of adding).

2.4.3 Use of Unknown Words

It is not unusual for students to confuse words they don't know when they are responding to an item. For instance, in a problem about buying snacks from a vending machine (Figure 2.22), a student may discuss buying vending machines rather than buying snacks. In figure 2.23 the student uses the word *"machine"* instead of *"snack"*.

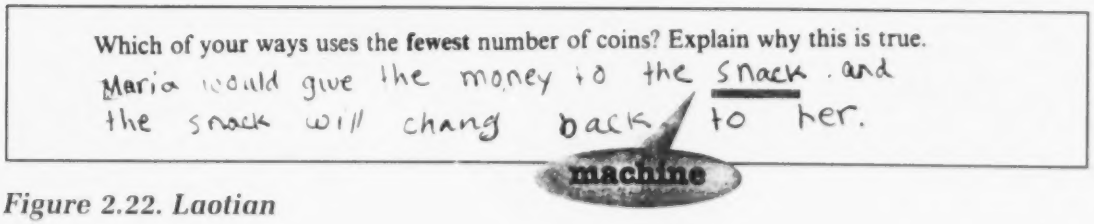

Which of your ways uses the fewest number of coins? Explain why this is true.
Maria would give the money to the snack and
the snack will chang back to her.
machine

Figure 2.22. Laotian

The Vending Machine

Maria wants to buy a 75-cent snack from a vending machine. The machine takes only nickels, dimes, and quarters. Maria has 7 nickels, 5 dimes, and 2 quarters.

Part 1

Show all of the different ways she could pay for the snack. You may use words, diagrams, or charts.

she have 5 dimes to buy a machine so she one machine whith 5$ and it so she could not ajy one.

Figure 2.23. Haitian Creole

2.4.4 Novice Sentence and Paragraph Structures

Other features of the second language development process include developmentally immature sentence and paragraph structures (e.g., chopped sentences) and little variation in sentence structure among sentences.

"I go market buy vegtabls of 5"

2.4.5 Limited Use of Language

Young learners of English as a first language use pictures and drawings as texts. This is a typical stage in their literacy development. Similarly, second language learners rely heavily on numbers, charts, and pictures so that there is a minimum use of language (Masuda, 1996). Again, depending on what is being measured, this may or may not be an acceptable response.

3.0 EFFECTS OF ENGLISH LANGUAGE DEVELOPMENT ON UNDERSTANDING AND INTERPRETING ASSESSMENT ITEMS

3.1 Cultural Influences

It is important to realize that students' responses to assessment items depend on their interpretation and understanding of the items. This interpretation and understanding is in part influenced by the student's language background and cultural reference. This is true for all students, and because of the open-ended nature of performance assessment items, a variety of interpretations is often possible, and appropriate, for the same item. However, ELL students' cultural backgrounds are often very different from the traditional American "educational culture" used in assessment, and may result in these students interpreting items in an unexpected fashion. There are a number of ways the misunderstandings described below might be handled in scoring responses, depending on how critical the misuse is to the content being assessed.

3.1.1 Misunderstanding the Meaning of Words

In figures 3.1 and 3.2, the students have interpreted "fewest" as smallest value.

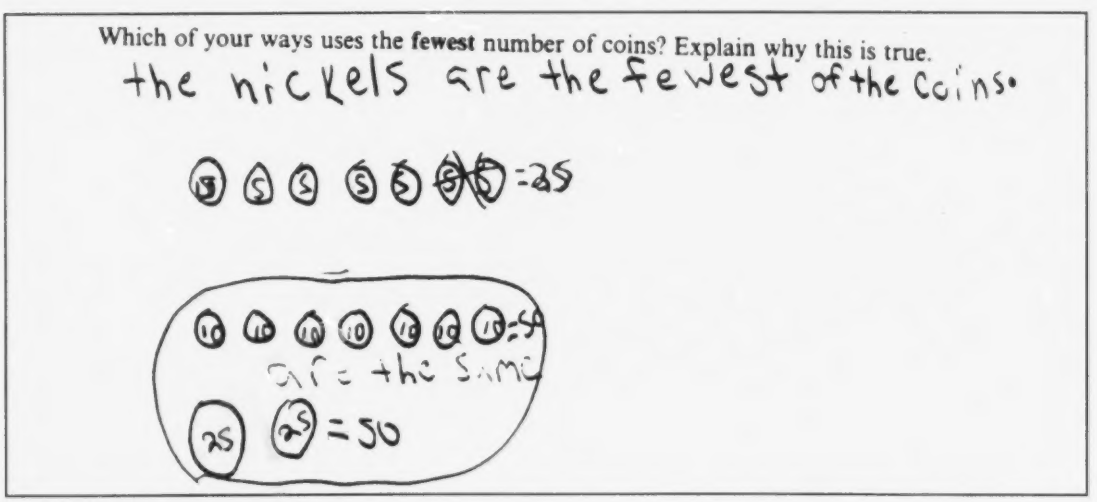


Figure 3.1. Spanish

Figure 3.2. Spanish

In figure 3.3, the student has interpreted "fewest" to be smallest in size.

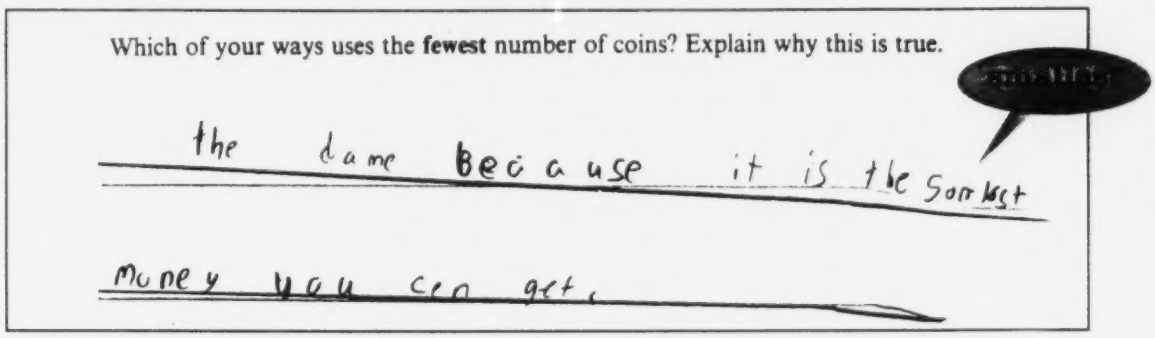

Which of your ways uses the fewest number of coins? Explain why this is true.
the dame because it is the Sorrest
Money you can get.

Figure 3.3. Spanish

3.1.2 Misunderstanding Systems and Symbols

Misunderstandings based on different systems of number, measurement, writing, and mathematics conventions, and different use of symbols could affect the interpretation of the requirements of the item. Examples of the different systems and symbols are explained in Section 2.2, Cultural Influences.

3.1.3 Misunderstanding Because of Differences in What Is Valued and Experienced

Students read an item based on their values and experiences, and the values and experiences prevalent in their culture. Since the values and experiences of the students' native culture and the U.S. culture may be dissimilar, there is room for misinterpretation.

Culture can affect a student's interpretation of an item in a variety of contexts. For example, an assessment item that asks students to create a fair race may elicit unexpected responses from some students. Whereas the creators of the item expect students to create a race-course in which all of the contestants have to run equal distances, some students may interpret fair to mean that all contestants have an equal chance of winning—this may be especially true in cultures that do not emphasize competition. As a result, these students may create a race course in which the slower contestants will run shorter distances. On the basis of their interpretation of a race and the notion of fairness, this is a valid response.

4.0 ISSUES RELATED TO THE CONSTRUCTION OF ACCURATE ASSESSMENTS FOR ELL STUDENTS

4.1 Validity Considerations

It appears that most of today's problems in accurately assessing academic achievement in English language learners are related to validity (Kopriva and Lowery, 1994). That is, developers assume that ELL scores on tests and their items are functions of the intended ability. Even the bias analyses assume that the responses are measuring the intended ability, once the influence of item response theory (IRT) parameters have been taken into consideration. This appears to be especially true if the analyses do not detect significant differences in item response curves, or if the results reflect what we know or what we think we know; for instance, if we assume that socioeconomic status is primarily responsible for affecting the response curves. Steps to validate the assumption of ability do not appear to have been done extensively by test developers.

4.1.1 Item and Test Development

The item formats common within performance assessment systems often require more reading, and more communication of knowledge, generally through writing. Further, as discussed above, the additional element of scoring the responses of English language learners can introduce significant error into the measurement of subject matter knowledge. For instance, the development and use of scoring criteria and guides, and the training and monitoring of scorers are intervention points which can provide situations that inversely affect the academic evaluation of ELL students, particularly because most scorers, to date, are monolingual native English speakers.

In attending to the potential problems inherent within traditional tests and tests that use a broader range of item formats, we need to be alert to challenges posed within the portions of the assessments that are presented to the students, that is:

- The directions
- The items and prompts
- Contextual passages (e.g., passages that introduce or explain a science experiment)

Other intervention points include recognizing and minimizing the unintended effects possible in:

- Administration procedures
- The composition of test forms
- Samples used in pilots/tryouts of items

- The equating procedures
- Procedures around mapping scores onto performance levels and standards
- The types and designs of technical studies
- The interpretation of and use of results
- Reporting

In their directions, items, and response criteria, large-scale assessments generally mirror the elements associated with the average literacy expectations of native speakers in a given grade. Unless the test is assessing reading or writing, it seems reasonable that the breadth and depth of literacy sophistication in tests be kept to a minimum, with the exception of certain vocabulary that is germane to a subject area and should be learned as part of learning the concepts the vocabulary represents. This includes mathematical concepts.

4.1.2 Language Complexity

The Council suggests the following guidelines as ways of dealing with the issue of language complexity for ELLs in assessments of subject areas other than language arts. These would also be useful within the assessment of language arts, when the particular skills noted below are not being specifically evaluated. In items, test directions, and contextual passages:

- Paraphrasing words or ideas should be recognized as such, and either not used in achievement assessments taken by ELLs or used with all the words in parentheses. This means using the same words to refer to the same phenomena, concept, or person, place, thing, action, or modifier rather than using a variety of words. While repetition may be considered bad writing by most teachers, it is necessary to present test questions that use a limited range of variant terms and constructions that make question comprehension easier for second language learners (Masuda, 1996).
- The range of organizational structures in sentences and paragraphs needs to be restrained. Again, this might be good writing, but it is common for ELL students to not recognize alternative ways of communicating the same thought, or to be confused by different approaches to discussions of an issue, or to a contextual explanation. A restricted, straightforward, consistent, and common (present or past) tense approach is best, and will lead to more accurate measurement of the subject under consideration.

4.1.3 Primary Language Assessment

We should not assume that ELL students have been taught the skills necessary to adequately take a primary language achievement test, as they are currently written. Such tests now are generally written in a style that assumes a sophisticated level of literacy in the primary language, consistent with the grade level literacy expectations of English speakers. Even though students may speak their native languge fluently, they are often not literate in the primary language. If primary language assessments are to be developed and used, it makes sense that at least some forms of the test be edited to reduce language complexity.

4.1.4 Item Overload

Item load is defined here to mean the amount of work required to successfully complete the demands of a test item, prompt, or task. Overload frequently can occur for English language learners because of the increased work they face in having to decode the language, in addition to reading the items and text for subject matter intent. This includes the intent of the distracters in close-ended items. Increased work requires additional time. It also significantly raises the probability of having fatigue and frustration affecting responses.

To address this problem, it is recommended that:

- Test developers specify additional administration and response time for ELL students.
- Test booklets be structured to ensure enough breaks for students. For instance, if a history assessment is generally designed to go over two class periods with a break in between, then it might be prudent to structure some forms with the same number and types of items per form but recommend that the test be completed with two breaks.

4.2 Recommendations

1. Test developers need to answer the question: Are we testing mathematics or language/mathematics? To get the highest score possible, should students be required to use English or their native language fluently to explain themselves? The National Council of Teachers of Mathematics (NCTM) standards, as well as many state standards, emphasize a student's ability to communicate about mathematics as an important component of knowing and demonstrating mathematics. Does it only mean communicating through writing? How about communication through algorithms? How should this be handled for all students, including ELL students?

2. Items should be reviewed specifically for linguistic and cultural inconsistencies. This review should include field testing where students and teachers are specifically asked to comment on words or phrases, and not just be a review by experts familiar with particular cultures.

3. Currently, most bias review panels focus on eradicating stereotypes in items and tests. The charge of review panels should be expanded to include a discussion and review of the type of skills that would be required to successfully complete complex performance items. This review should be supplemented with student work, field test data, and evaluations. Recommendations should be made about item requirements.

4. Flexibility in both administration and response formats can noticeably affect the accuracy of ELL scores. LaCelle-Peterson and Rivera (1994) emphasize that some students initially develop a higher level of listening and speaking proficiency in English as compared to their reading and writing English proficiency levels. This appears to be especially true when the students are not literate in their primary language (Wong-Fillmore, 1994). Therefore, the reading and writing formats traditionally used in large-scale assessments may not be an effective way to measure subject matter achievement. While Kopriva and Lowrey (1994) found that cueing regularly occurs when teachers are allowed to read the test aloud to students, it is recommended that some kind of standardized read-aloud approach, perhaps using cassette tapes, be available to students to use in appropriate situations. It is also recommended that tapes or a computer voice recognition program be used to enable students to respond orally to performance items.

5. To ensure that assessments are actually measuring the types of skills and knowledge inherent in the items rather than coping skills needed to deal with novel approaches, students must have similar experiences in their classrooms. That is, they must have ongoing classroom opportunities that require them to use content by demonstrating critical thinking skills, and that require them to verbalize and write about these experiences in the language of the test.

6. Repeated classroom experiences with the types of items and tests that are being used in large-scale situations are also important, if the assessments are going to accurately reflect what students know. This includes (a) the setting of the tests, e.g. on-demand or completed over days, with or without feedback from others; (b) item formats, i.e., open or closed, short answer, or extended response, and (c) the response requirements, i.e. written numeric, or graphic, and first or finished draft.

5.0 SUMMARY

To date, little work has been done to ensure that LEP students are accurately assessed on large-scale tests (Olsen and Goldstein, 1997). The purpose of this guide is to help scorers in high-volume situations be able to effectively evaluate the open-ended responses of this population. Section 1 presents a brief overview about CCSSO's LEP Project and some background information about the nature of the development of language proficiency. Section 2 provides guidance about linguistic issues in the students' native language, developmental benchmarks associated with learning English, and cultural influences—all of which affect the written responses of LEP students. This section is illustrated with numerous examples and provides specific recommendations about how to accurately read the responses.

Sections 3 and 4 place the scoring of items in the larger context of test development and implementation. Section 3 briefly discusses the effects of English language development on how ELL students understand and interpret some assessment items, and Section 4 provides an overview of how tests might be constructed and administered to ensure the validity of the tests for these students. Recommendations throughout Section 4 highlight where and how to improve large-scale testing for English language learners.

The Guide is designed so that it can be efficiently integrated into the regular training scorers receive when they are preparing to score test items. Used with the accompanying training guide, it has been found to be a useful tool for scoring LEP student work in large-volume settings (Kopriva, 1997). As more and more English language learners are included in large-scale tests, this guide will help ensure that their work is evaluated reliably and responsibly, so that the real progress toward meeting challenging standards can become a reality for all.

Rebecca Kopriva is currently an independent consultant. Dr. Kopriva worked previously as Director of Assessment in the Delaware Department of Public Instruction, and while a Professor at California State University, as a Coordinator of assessment studies on equity for the California state test, CLAS.

Sharon Saez is an Education Program Specialist for the U.S. Department of Education's Goals 2000 Program. She was formerly with the Delaware Department of Public Instruction as lead developer of the Delaware State Testing Program, and worked previously as Title VII Coordinator.

Appendix A

SCASS LEP PROJECT

TRAINER's MANUAL DEVELOPMENT

COMMITTEE

California

Holga Davis
Sacramento

David Delgado
San Bernadino

Florida

Ana Lopez
Miami

Yuria Alban
Miami

Lourdes Rovira
Miami

Texas

Ann Mitchell
Austin

Raul Ramon
Corpus Christi

Daniel Garcia
Weslaco

Delaware

Lynn Smail
Wilmington

Victoria Antinguiani
Wilmington

Connecticut

George P. George
Hartford

CCSSO Staff

Julia Lara
Director, Services to Language Minority Students

Edward Roeber
Director, CCSSO Student Assessment Programs

Reviewers

Laura Ayala
Texas Education Agency

Michael Ford
University of California, Berkeley

Walter Masuda
University of California, Berkeley

Dominick Perissini
University of Colorado, Boulder
Fomerly with NCMRSE
University of Wisconsin, Madison

Mary Ramirez
School District of Philadelphia Public Schools

Appendix B

References

Collier, V.P. (1987). *Age and rate of acquisition for academic purposes.* TESOL Quarterly, 21. 617-641.

Ford, M. (1976). Personal Communications.

Fradd, S., and Weismantel, M.J. (1989). *Meeting the needs of culturally and linguistically different students: A handbook for educators.* New York: Little, Brown and Company.

Khisty, L. (1995). Making inequalities: Issues of language and meaning in mathematics teaching with Hispanic students. pp. 279-297. In Secada, W., Fennema, E., and Adajian, L.B. (Eds.), *New directions for equity in mathematics education.* Cambridge, UK: Cambridge University Press.

Kopriva, R., and Lowrey, K. (1994). *Investigations of language sensitive modifications in a pilot study of CLAS, the California Assessment System.* Sacramento: California Department of Education, California Learning Assessment System Unit.

Kopriva, R. (1997). *Piloting the Large-Scale Assessment Training Guide for Scoring LEP Students.* Unpublished Report.

La Celle-Peterson, M. and Rivera, C. (1994). *Is it real for all kids? A framework for equitable assessment policies for English language learners.* Harvard Educational Review, 641(1), 55-75.

Lipski, J.M. (1985). *Linguistics aspects of Spanish-English language switching.* Tempe, Arizona: Center for Latin American Studies.

Masuda, W. (1976). Personal Communications.

Olson, J.F., and Goldstein A.A. (1997). *The Inclusion of Students with Disabilities and Limited English Proficient Students in Large-Scale Assessments.* Washington, DC: National Center for Education Statistics.

Valette, R.M. (1997). *Modern language testing: A handbook.* New York: Harcourt Brace Jovanovich.

Wong-Fillmore, L. (1994). The role and function of formulaic speech in conversation. In Grimshaw and D. Jennes, (Eds.), *What's going on here? Complementary studies of professional talk.* New York: Ablex Publishers.

Appendix C

Glossary

Access. Rights and means to approach or engage in with understanding. Assessments provide for equal access when they include items that are shown to be equally appropriate for all students, allow multiple approaches and strategies, and accept multiple justifiable responses.

Assessment. The process of gathering evidence about a student's knowledge of, ability to use, and disposition toward mathematics and of making inferences from that evidence for a variety of purposes. *Assessment* is a term that has often been used interchangeably with the terms *testing*, *measurement*, and *evaluation*, or to distinguish between student assessment and program evaluation. In this document, *assessment* is used as defined above to emphasize understanding and description of both qualitative and quantitative evidence in making judgments and decisions.

Assessment Instruments. A set of assessment items and/or activities for a particular academic area that are tied to content and student performance standards and used for assessing student academic performance and achievement.

Assessment Items. Test questions or activities based on content used to determine students' achievement of performance standards.

Benchmarks. Descriptions of student performance at various developmental levels that contribute to the achievement of performance standards.

Bias. The systematic undermeasurement or overmeasurement of a student's "true" skill.

Code Switching. The alternate use of two languages. Language selection depends on the context.

Content Standards. Statements about what it is that students must know and be able to do in various disciplines such as English language arts, mathematics, science, and social studies.

Curriculum. (1) An educational program that may include a program of studies, a program of activities, and a program of guidance; (2) an operational plan for instruction that details what students need to know, how students are to achieve the identified curricular goals, what teachers are to do to help students develop their knowledge, and the context in which learning and teaching occur.

Curriculum Framework. A comprehensive document outlining the broad goals and content standards of an entire system of education, while leaving the local district the freedom to develop a specific program to address the framework.

Embedded Assessments. Assessments that are developed, integrated, and used within an instructional context.

Equitable Assessment. The degree to which the process of gathering evidence has provided opportunities equally appropriate for each student to demonstrate the valued thinking processes, knowledge, and skills that he or she has developed. Equitable assessment is not achieved by creating the same conditions for all students but rather by creating conditions that are appropriate to the same extent for each student.

Equity. Fairness; a matter of equal opportunity; providing for each student the opportunity he or she needs to succeed educationally. For example, tests and assessments should not systematically penalize an individual because of gender, race, or cultural background. Likewise, differences in educational programs to address individual student needs should not systematically offer some students less rich educational experiences.

Generalizations. Inferences or conclusions from many particulars of the evidence in hand, supported by a theory of the relationships between the particulars and the more general inferences or conclusions.

Home Language. The primary language learned by the child, usually the language spoken at home.

Inferences. Conclusions or assertions derived from evidence; deductions.

Integration. The idea that more than one discipline can and should be taught or assessed at the same time, or that behaviors of thought (e.g., problem solving and referencing) are not the exclusive domain of any one discipline. Integrated instruction connects subject areas in ways that reflect the real world. For example, a student's writing in his or her science or social studies class can serve as a sample of writing for assessment purposes. The literature studied in English class can reflect and be taught in concert with history or world cultures units or with scientific concepts and theories under discussion in science. Key scientific discoveries can be studied in concert with their effect on the history of a nation or people.

Item. A single test question or problem.

Judgments. Authoritative estimates or opinions of quality, value, and other features, formed by distinguishing the relations among multiple sources of sound and reasonable evidence; formal decisions.

Language Dominance. The language in which a bilingual person has the greatest command. Dominance in one language or another can vary depending on the context or situation.

On-Demand Assessments. Assessments are administered at a specific time under standardized testing conditions.

Open-Ended Questions. Items that ask student to formulate their own response, which typically take anywhere from two to twenty minutes to complete. Open-ended problems engage students in interesting situations and allow students at many levels of understanding to begin working on the problems, make their own assumptions, develop creative responses, and effectively communicate their solutions.

Opportunity to Learn. The degree to which a student has been exposed to the learning experiences needed to meet high academic standards, which is largely a function of the capacity and performance of the courses and schools the student has attended. Equitable opportunities to learn consist of equal chances for learning, with equally appropriate, favorable, or advantageous combinations of circumstances; i.e., opportunities to learn are equitable when they are responsive to the same extent to each student's needs.

Performance. The carrying out or bringing to completion of an activity or production of some significance, which displays a student's knowledge and judgment while engaged in the task.

Performance Criterion or Standard. A statement of expected performance quality that can be used to make judgments about performances that are central to the curriculum. A set of performance criteria, or standards, includes the nature of the evidence required and the quality of performance expected to demonstrate that a curriculum or content standard has been achieved. These statements often describe performances at one level, such as either adequate or exemplary, but may also describe a range of quality levels.

Performance Assessments. Assessments that require students to demonstrate what they know and can do according to established education standards. Performance events are assessment activities that require students not only to choose the best answer, but to explain the reasoning behind the answer or solve a problem either individually or with a group of students.

Reliability. The degree to which multiple scores are consistent with one another.

Rubric. A set of authoritative rules to give direction to the scoring of assessment items or activities. To be useful, a scoring rubric must be derived from careful analysis of existing performances of varying quality. An **item-specific rubric** describes levels of performance for a particular complex performance item and guides the scoring of that item consistent with relevant performance standards. (A item-specific rubric is more specific than a performance standard and can apply a performance standard to a particular context found in a performance item.) A **general rubric** is an outline for creating item-specific rubrics, or for guiding expert judgment, where item-specific scoring rules are internal to the scorer.

Scoring. Discriminating among performances according to differing levels of quality and assigning a descriptive label or number to the performance. In **holistic scoring,** the entire performance as a whole is considered, and one label or number is assigned. In **analytic scoring,** separate scores are assigned to fundamentally different dimensions of the performance.

Standard-Referenced Assessment. An assessment that compares the quality of performances to relevant performance criteria or standards to make a determination of the degree to which the standards have been attained or to describe progress toward the attainment of the standards.

Technical. Relating to formal, psychometric determinations of the quality of scores.

Valid. Justifiable, well grounded, sound; producing the desired results, efficacious; incontestable.

Valid Inference. Justifiable assertions and conclusions that lead to and support desirable results. Justification is made primarily on the quality of the evidence and its adequacy for the intended purposes and their consequences.

Validity. Refers to whether the assessment instrument measures what it purports to measure.